

# Computer schneiden Grimassen

Animierte Figuren etwa in Filmen und in Computerspielen wirken heute oft lebensecht. Denn sie werden mit aufwendig erzeugten, dreidimensionalen Modellen von Körpern und Gesichtern erschaffen. **Christian Theobalt** und seine Mitarbeiter am **Max-Planck-Institut für Informatik** in Saarbrücken erleichtern es Grafikern deutlich, solche Modelle zu erstellen. Sie ermöglichen dadurch Anwendungen, die bislang undenkbar waren.

TEXT **TIM SCHRÖDER**

**D**er Schauspieler Tom Cruise ist ein cooler Typ. Im Film *Minority Report* glänzt er als hartgesottener Kämpfer für das Gute. Er spielt John Anderton, einen Polizisten im Jahr 2054 – der, so gehört es sich für Science-Fiction, von jeder Menge Hightech umgeben ist. Sieht man sich den Film heute, nur 14 Jahre nach der Kinopremiere an, wirkt die Zukunftsvision streckenweise gar nicht mehr so fortschrittlich. In einer damals noch futuristisch anmutenden Szene lässt der Schauspieler auf einer Glaswand, die als Monitor dient, per Handbewegung Bildschirmfenster aufleuchten. Per Fingerzeig zieht und staucht er die Bilder, mit einer flotten Wischbewegung zaubert er sie fort. Das sieht chic aus, allerdings trägt er dabei einen schwarzen Handschuh mit Leuchtpunkten – einen Datenhandschuh, den *data glove*. Und dafür haben Wissenschaftler heute allenfalls noch ein müdes Lächeln übrig.

Auf das Prinzip, nach dem Tom Cruise alias John Anderton per Datenhandschuh einen Computer steuert, setzt die Filmindustrie seit rund 20 Jahren auch bei der Produktion neuer Streifen. Sie filmt eine Person, die einen ganzen Anzug mit Markierungen, sogenannten Markern, trägt, damit ein Computer anhand der Markierungen die Position von Kopf, Rumpf, Armen und Beinen nachverfolgen – tracken – kann. Damit lässt sich dann die Bewegung in eine Filmszene oder eine Computerwelt übertragen, zum Beispiel, um eine fiktive Figur menschenecht zu animieren. Doch diese Art des Trackings ist recht umständlich. Der Schauspieler muss sich in Handschuhe und Anzug zwängen.

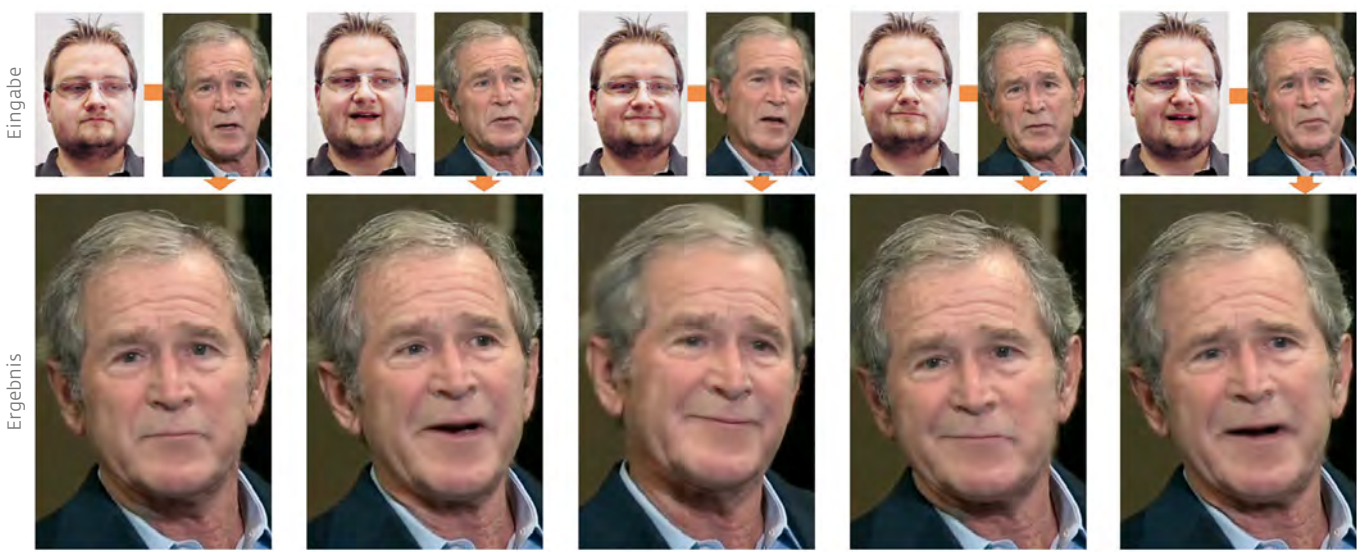
Dass das Tracking mit Markern allmählich aus der Mode kommt, ist unter anderem Christian Theobalt und seinen Mitarbeitern am Max-Planck-Institut für Informatik in Saarbrücken zu verdanken. Der Informatiker ist dort Leiter der Arbeitsgruppe „Grafik, Sehen

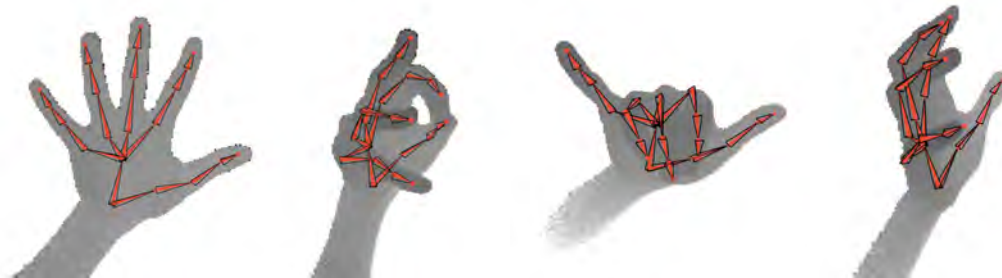
& Video“. Die Wissenschaftler arbeiten an der Grenze zwischen Computergrafik und Bilderkennung. Das große Ziel Theobalts ist es, dem Computer beizubringen, die sich bewegende Realität mit möglichst wenig Aufwand und in Höchstgeschwindigkeit zu analysieren und daraus detailgenaue und vor allem dreidimensionale virtuelle Modelle zu erzeugen, welche sowohl die Form als auch die Reflexionseigenschaften einer Figur sowie die Beleuchtung einer Szene richtig erfassen.

## **BEDARF FÜR DIE METHODE HABEN NICHT NUR FILMEMACHER**

Seine Arbeitsgruppe hat hier unter anderem zwei Forschungsschwerpunkte: erstens die Erkennung von Bewegungen und zweitens die Analyse von Gesichtern. Bedarf für diese Methoden haben nicht nur Filmemacher, die mit den Modellen etwa bei der Mimik eines Schauspielers nachhelfen oder Fantasiefiguren

Manipulierte Mimik: Saarbrücker Informatiker übertragen den Gesichtsausdruck einer Person in einem Quellvideo auf eine Person in einem Zielvideo. Die grauen und braunen Pfeile geben den Ablauf wieder. In den ersten beiden Reihen sind die beweglichen Gesichtsmodelle im Quell- und Zielvideo als Schachbrettmuster über die Gesichter gelegt. In der vierten und sechsten Reihe von oben ist die Originalmimik von Quelle und Ziel zu sehen.





Richtige Gesten: Franziska Müller (rechte Seite) hat eine Software entwickelt, die sogar subtile Bewegungen einer Hand mit relativ geringem Rechenaufwand verfolgt und die Gesten in einem 3D-Modell rekonstruiert. Um sicherzustellen, dass die wiedergegebenen Haltungen anatomisch möglich sind, verwendet das Programm das Modell eines Handskeletts, das in Rot dargestellt ist (diese Seite).

animieren können. Die Entwickler von Computerspielen könnten auf ähnliche Weise Avataren zu natürlichen Bewegungen und Gesichtsausdrücken verhelten. Und auch für Roboter könnten die Techniken der Szenenanalyse nützlich sein, damit diese sich von ihrer Umgebung ein genaues Bild machen können. Vor allem müssen sie Menschen im Hintergrund von relevanten Objekten im Vordergrund unterscheiden können. Ein weiteres Anwendungsgebiet ist die *augmented reality*, „erweiterte Realität“, also die Vermischung von Realität und virtueller Welt, bei der man mit Gesten etwa virtuelle Objekte bewegen kann.

Allein der Film- und der Spielebranche dürfte Theobalts Forschung weltweit zukünftig Tausende von Personenjahren an Arbeit ersparen, denn die Animation von Filmen und Videosequenzen ist heute nicht nur wegen des Trackings mit Markern eine Heidenarbeit. Grafiker müssen Aufnahmen einer realen Person in ein mathematisches Modell übersetzen, um daraus dann Figuren zu generieren und diese in eine Computerspiel- oder Filmsequenz zu übertragen.

Heute ist dafür noch viel Handarbeit notwendig. Damit zum Beispiel das künstliche Gesicht eines Schauspielers, das man in eine besonders spektakuläre Filmsequenz hineinkopiert, echt lächelt oder die Stirn kräuselt, müssen die Computergrafiker viele Stunden Feinarbeit investieren.

Was den ersten Teil seiner Forschung betrifft – die Bewegungsanalyse, das *motion capturing* –, sind heute noch ei-

nige Hürden zu überwinden, sagt Theobalt. Um zum Beispiel den gruseligen Gollum in den *Hobbit*- und *Herr-der-Ringe*-Filmen auftreten zu lassen, haben die Filmemacher den Körper und die Bewegungen eines Schauspielers in ein Modell übertragen. Die Gestalt und vor allem das Gesicht haben sie dann an ihre Vorstellungen des Fantasiewesens angepasst, die Figur mit der Gollum-Haut texturiert und sie schließlich in eine am Computer erzeugte Höhlenwelt hineinkopiert.

### BEWEGUNGEN VON HÄNDEN ZUVERLÄSSIG ERKENNEN

Damit die Bewegung des Schauspielers vom Computer in ein Modell umgerechnet werden kann, muss der Schauspieler in speziellen Studios perfekt ausgeleuchtet und von mehreren Kameras aus verschiedenen Richtungen aufgenommen werden; nur so werden auch verdeckte Körperteile erfasst. Und natürlich sollten die Aufnahmestudios bis auf den Schauspieler relativ leer sein, weil Gegenstände im Raum den Computer bei der Analyse der Szene verwirren würden.

„Viel einfacher wäre es, wenn man Personen unter freiem Himmel, in einer ganz normalen Umgebung und bei wechselnden Lichtverhältnissen aufnehmen könnte“, sagt Theobalt. „Und am besten mit nur einer Kamera, um den Aufwand möglichst gering zu halten – genau das ist unser Ziel.“ Bis vor Kurzem war das noch undenkbar. Doch

Schritt für Schritt kommen Christian Theobalt und seine zehn Mitarbeiter diesem Ziel näher.

Die Doktorandin Franziska Müller zum Beispiel will dem Computer zusammen mit ihrem Kollegen Srinath Sridhar helfen, die Bewegungen von Händen mit einer Kamera zuverlässiger zu erkennen und in ein dreidimensionales Modell zu übertragen. Das ist besonders schwierig, weil bei den raschen Bewegungen einer Hand und ihrer Finger immer wieder Teile aus dem Blickfeld der Kamera geraten. Den Fingern genau folgen zu können ist aber unter anderem wichtig, um Geräte etwa in der *augmented reality* mit Gesten zu bedienen oder Robotern bestimmte Bewegungen beizubringen. „Dazu muss der Computer die Geste der Hand richtig deuten können“, sagt Franziska Müller.

Das, was Tom Cruise in *Minority Report* vollführt hat, hat Müller dabei bereits weit hinter sich gelassen. Cruise arrangierte mit groben Handbewegungen die Bilder auf einem Monitor um. Franziska Müllers Computer aber kann feinste Greifbewegungen erkennen. Dazu schaltet sie eine kleine Kamera an, die auf ihrem Computerbildschirm sitzt und mithilfe von Laserstrahlen auch die dreidimensionale Gestalt eines Objekts misst. Prompt erscheint auf dem sonst weißen Bildschirm eine künstliche Hand, die jeder ihrer Fingerbewegungen folgt, zum Beispiel wenn Franziska Müller Daumen und Zeigefinger gegeneinanderdrückt oder ihre Hand öffnet und schließt.





Wie so oft steckt der Teufel dabei im Detail. Denn der Computer muss die Position der Finger ständig neu berechnen, und zwar in Sekundenbruchteilen, weil das Bild auf dem Computer sonst ruckeln und stocken würde. „Das ist nur mit mathematischen Verfahren möglich, die die Menge der Bilddaten stark reduzieren und trotzdem die Position richtig berechnen können“, sagt Franziska Müller.

Konkret geht es dabei um die mathematische Analyse von Entfernungsdaten. So misst die kleine Tiefenkamera auf Franziska Müllers Monitor für jeden Bildpunkt die Laufzeit des Lichts – etwa zu einer Fingerkuppe oder zum Handballen und zurück.

Die Doktorandin schaltet ein anderes Programm ein, das sie von einer Softwarefirma gekauft hat, die bereits ein Programm zur Echtzeit-Handmessung anbietet. Das Ergebnis ist enttäuschend: Als die Forscherin ihre Hand schnell bewegt, kommt das Programm nicht mehr hinterher. Das Modell der Hand auf dem Monitor verliert plötzlich Finger, einmal taucht ein Finger an einer falschen Stelle auf. Ganz schlimm wird es, als ein Finger einen anderen verdeckt. Die Hand auf dem

Bildschirm zerfließt stellenweise. Die gekaufte Software hat Probleme, die Bilddaten der Kamera korrekt in das Modell zu übersetzen.

Der Grund: Herkömmliche Programme schaffen den enormen Rechenaufwand nicht, um anhand von Bildern aus einer Kameraperspektive stabil die Bewegungen in drei Dimensionen zu errechnen. Müller setzt deshalb ein anderes Verfahren ein als das kommerzielle Programm. Ihre Software ordnet die Messwerte für einzelne Pixel so, dass benachbarte Pixel, die dieselbe Entfernung zur Kamera haben, als mathematische Wolken, als Gauß-Wolken, dargestellt werden. So lässt sich die Zahl der Punkte deutlich reduzieren. Die Rechenzeit verkürzt sich. Deshalb kann Müllers Programm auch dann Schritt halten, wenn die Forscherin ihre Hand schnell bewegt.

Die Berechnungen aus den Entfernungsmessungen gleicht Müllers Software mit einem Skelettmodell ab, das im Programm bereits hinterlegt ist. Es gibt dem Computer Anhaltspunkte, welche Haltung der Hand und welche Fingerstellungen möglich sind.

Zudem nutzt Müller ein maschinelles Lernverfahren, das aufgrund von

Wahrscheinlichkeiten in Sekundenbruchteilen abschätzt, zu welchem Teil der Hand ein Pixel gehört. Zu diesem Zweck hat Franziska Müller den Computer mit Trainingsdaten gefüttert: Er hat gelernt, wie eine Hand aussehen kann, wenn sie gedreht oder bewegt wird. Außerdem hat Müller in ihr Programm noch eine Art Fehlerabschätzung eingebaut, die Haltungen ausschließt, die gemäß der Handanatomie keinen Sinn ergeben.

„Dank Franziskas Arbeit können wir jetzt auch feingliedrige Bewegungen auflösen, zum Beispiel wenn man Daumen und Zeigefingerkuppe aneinanderreibt“, sagt Christian Theobalt. „Mit herkömmlichen Programmen ist das nicht möglich.“

Die Hand ist natürlich nicht alles. In vielen Fällen muss die Bewegung eines ganzen Körpers erfasst werden. Auch dafür nutzen die Saarbrücker ein Skelettmodell, das in den vergangenen Jahren in Theobalts Arbeitsgruppe entwickelt wurde und ihrer Software gewissermaßen anatomisches Wissen gibt.

„Damit lösen wir uns von der klassischen Bewegungsanalyse“, sagt Theobalt. Normalerweise orientiert sich der Computer in einer Folge von Bildern an

» Theobalt will die Gesichtserkennung auf ähnliche Weise vereinfachen wie die Bewegungsanalyse. Sein Team arbeitet daran, Filmaufnahmen, die eine einzige Kamera gemacht hat, in ein 3D-Gesichtsmodell zu übertragen.

charakteristischen Strukturen, nicht nur an Markerpunkten, sondern auch Bildregionen, die ein ähnliches Aussehen haben. „Wir nennen dieses Verfahren Korrespondenzfindung. Der Computer versucht damit, einem Gegenstand zu folgen, der sich in einer Bildsequenz langsam weiterbewegt.“ Das Problem: Bei wechselnder Beleuchtung erzeugen diese Verfahren deutlich mehr Fehler, weil die korrespondierenden Bildpunkte ständig ihre Helligkeit verändern.

Theobalts Team hat die Bewegungsanalyse nicht nur unabhängiger von der Umgebung gemacht, sondern die Zahl der dafür nötigen Kameras von mehr als acht auf drei reduziert. Zu diesem Zweck setzt Theobalt auch hier maschinelle Lernverfahren ein. So können die Forscher wettmachen, dass der Computer verdeckte Körperteile bei nur wenigen Kameras und wechselnder Beleuchtung leicht für kurze Zeit aus den Augen verlieren kann. Sie trainieren das Programm für maschinelles Lernen mit Bildern verschiedener Posen, sodass es lernt, die Körperteile sicher zu identifizieren.

Dieser kombinierte Ansatz macht die Bewegungsanalyse von Theobalts Gruppe besonders leistungsfähig. Es sei das erste Verfahren, sagt der Wissenschaftler, das die Bewegung des kompletten Skeletts unter diesen Bedingungen in 3D schnell und robust

messen könne – also nicht nur im sorgfältig ausgeleuchteten Studio, sondern auch draußen in einer beliebigen Umgebung, unter ständig wechselnden Lichtverhältnissen.

### THECAPTURE IST SPEZIALISIERT AUF BEWEGUNGSANALYSE

Inzwischen haben ehemalige Doktoranden und Postdoktoranden Theobalts die Firma TheCapture gegründet, die sich auf Bewegungsanalyse mithilfe des Skelettmodells spezialisiert hat. Das Unternehmen bietet eine Software an, die aus Videoaufnahmen einer oder weniger Kameras die Position und Bewegung der Gliedmaßen analysiert, auch in Echtzeit. „Die Software wird zur Analyse schneller Bewegungsabläufe bei Sportlern verwendet oder auch um die Körperhaltung von Menschen am Arbeitsplatz zu untersuchen“, erklärt Christian Theobalt.

Die Herausforderungen bei der Gesichtserkennung, dem zweiten Schwerpunkt der Saarbrücker Arbeitsgruppe, sind durchaus ähnlich. Um heute realistisch wirkende hochaufgelöste 3D-Gesichtsmodelle zu erzeugen, muss das Gesicht einer Person definiert ausgeleuchtet und von mehreren Kameras aufgenommen werden. Nur dann kann der Computer die dreidimensionale Form des Ge-

sichts errechnen und Fältchen sowie Reflexe der Haut sauber rekonstruieren. Um etwa das Gesicht eines Schauspielers in künstliche Welten zu versetzen, muss dieser viele verschiedene Gesichtsausdrücke einspielen: Er muss etwa lachen, böse gucken, die Augenbrauen heben. Eine Mimik, die nicht aufgenommen wurde, muss in einer Filmszene aufwendig in das Gesicht modelliert werden.

Gesichtserkennung spielt aber nicht nur in der Film- und Computerspiele-Industrie eine Rolle, sie ist auch für viele andere Anwendungen interessant. So kommt es bei neuartigen Müdigkeitswarnern im Auto darauf an, Gesichtszüge sicher zu deuten. Manche Firmen arbeiten außerdem an Verfahren, um Lippenbewegungen zu interpretieren. Damit ließe sich zum Beispiel die automatische Spracherkennung deutlich verbessern, da man nicht nur den Audiokanal nutzen könnte, sondern auch die Lippenbewegung im Videobild.

Theobalt will die Gesichtserkennung auf ähnliche Weise vereinfachen wie die Bewegungsanalyse. Und dabei dreidimensionale Modelle schaffen, die auch Gesichtsausdrücke wiedergeben können, wenn diese bei der Erzeugung des Modells nicht eingespielt wurden. Sein Team arbeitet daran, Filmaufnahmen, die eine einzige Kamera bei beliebiger Beleuchtung von Gesichtern ge-

Die Methode von Christian Theobalts Team analysiert auch die Bewegungen ganzer Körper mithilfe eines Skelettmodells – hier für einen Boxer gezeigt.







Die bewegte Realität zu analysieren und dabei nicht nur die Formen von Körpern und Gesichtern einzufangen, sondern auch die Reflexionseigenschaften und die Beleuchtung einer Szene – das ist die Mission von Michael Zollhöfer, Franziska Müller, Abhimitra Meka, Dushyant Mehta, Hyeonwoo Kim, Pablo Garrido und Christian Theobalt (von links).

macht hat, in ein 3D-Gesichtsmodell zu übertragen. Anders als herkömmliche Rechenverfahren arbeitet die Technik so schnell, dass das Modell der Mimik einer gefilmten Person folgen kann.

Um ein bewegtes Kunstgesicht in annehmbarer Zeit aus dem einfachen Videobild einer einzigen Kamera zu rekonstruieren, muss Theobalt einen völlig anderen Weg als bisherige Methoden gehen. Er nennt ihn *inverse rendering*, frei übersetzt: umgekehrte Wiedergabe. Der Begriff *rendering* kommt aus der Computergrafik und steht für die genaue Berechnung von korrekt ausgeleuchteten Bildern aus einem Modell der Szene. Im inversen Rendering dreht man das um, man errechnet das Modell der Beleuchtung, Reflektanz und Geometrie, das am besten das Aussehen und die Schattierungen im Bild erklären kann. Die Gesichtsrekonstruktion wird dadurch sehr robust gegenüber Szenenveränderungen und funktioniert unabhängig davon, ob die Sonne scheint oder der Himmel vor dem Fenster verhangen ist.

Der Trick: Statt wie in herkömmlichen Verfahren ein Gesicht mitsamt Falten, Schatten und Reflexen unter Studiobedingungen Pixel für Pixel zu analysieren, teilt Theobalts Team die Kon-

struktion des Modells in vier parallele Schritte auf: erstens die Erkennung der Gesichtsform; zweitens die Rekonstruktion, wie sich diese bei verschiedenen Gesichtsausdrücken verändert; drittens die Abschätzung der Reflexionseigenschaften der Gesichtsoberfläche, der sogenannten Reflektanz; und viertens die Schätzung der Beleuchtung im Raum.

### RASCHE VERÄNDERUNGEN FLIESSEND WIEDERGEBEN

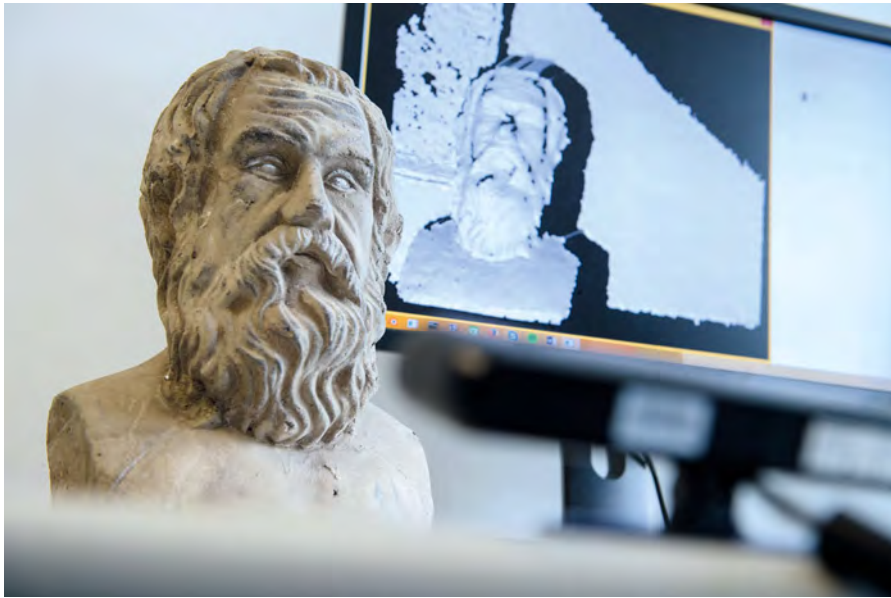
Bei der Erkennung der Gesichtsform und ihrer Veränderungen besteht die Herausforderung darin, aus dem zweidimensionalen Videosignal der Kamera räumliche Informationen zu gewinnen – die Position der hervorstehenden Nase oder der tiefer liegenden Augen beziehungsweise die Form des Mundes. „Wir überlagern die Aufnahme des Gesichts deshalb mit einem 3D-Gesichtsmodell, das vor einigen Jahren hier am Institut entwickelt wurde“ (siehe *MAXPLANCKFORSCHUNG* 3/2011), sagt Theobalt. „Dessen Stärke ist es, aus wenig Bildinformation ein räumliches Gesicht zu rekonstruieren.“

Separat von der Erkennung der Form wird beim inversen Rendering die Information über Licht und Schatten in

einer Szene beziehungsweise die Reflektanz analysiert. „Wir berechnen daraus zunächst, welche Lichtverhältnisse im Raum herrschen“, sagt Theobalt. Anschließend kann der Computer das Licht im Raum mit der Gesichtsform reagieren lassen. Damit kann er dann aus Schattierungen des Gesichts im Video wieder auf die 3D-Form schließen, inklusive feiner Details im Gesicht. In mehreren Schleifen gleicht der Computer das von ihm erzeugte Gesichtsmodell in Sekundenbruchteilen immer wieder mit dem tatsächlichen Videobild ab und passt es entsprechend an, bis es mit der Vorlage übereinstimmt. Das geht so schnell, dass das Modell selbst rasche Veränderungen der Gesichtszüge fließend wiedergeben kann.

Wie gut das inverse Rendering funktioniert, hat vor Kurzem Theobalts Postdoktorand Michael Zollhöfer zusammen mit Kollegen von den Universitäten in Stanford und Erlangen gezeigt. Die Forscher sorgten für Medienrummel, als es ihnen gelang, die Mimik eines Gesichts in Echtzeit auf ein anderes zu übertragen – *reenactment*, Nachstellung einer Mimik, heißt der Vorgang.

Zollhöfer zeigt, wie es geht. Er schaltet eine herkömmliche Kamera von der Größe eines Schokoriegels ein und



Rekonstruierte Antike: An einer Büste von Sokrates haben die Saarbrücker Informatiker die Stärke des inversen Renderings vorgeführt. Mit der Methode lassen sich Oberflächen detailliert erfassen, auch wenn diese mit kommerziellen Tiefenkameras aufgenommen werden, die normalerweise nur grob aufgelöste 3D-Modelle liefern.

nimmt damit sein Gesicht auf. Auf dem Monitor erscheint Zollhöfers Kopf, den der Computer zunächst mit einem Gitternetz überzieht. „Der Computer errechnet jetzt das Modell meines Gesichts, das dauert ein paar Sekunden“, erklärt Zollhöfer. Dann aber geht es schnell. Wie eine venezianische Karnevalsmaske erscheint jetzt auf einem zweiten Bildschirm der dreidimensional animierte Gesichtsschnitt Zollhöfers. Bewegt Zollhöfer jetzt seinen Mund, so folgt die Maske der Bewegung.

### EINIGE PRODUZENTEN HABEN BEREITS ANGEKLOPFT

Dann schaltet er auf einem zweiten Monitor ein Video hinzu, das Arnold Schwarzenegger in einem Interview zeigt. Auch von Schwarzeneggers Gesicht erzeugt die Software ein Modell, das gewissermaßen hinter Schwarzeneggers Videogesicht liegt. Dann der Knaller: Als Zollhöfer seinen Mund öffnet, öffnet auch Arnold den Mund. Zollhöfer rümpft die Nase, er grinst, zieht die Stirn kraus – und Arnolds Bild folgt brav jeder Bewegung. „Wie man sieht, wird mein Gesichtsausdruck in Echtzeit auf das Gesichtsmodell Arnold Schwarzeneggers übertragen“, erläutert Michael Zollhöfer.

Für die Filmindustrie bedeutet das, dass eine natürlich wirkende Mimik direkt von einer Person in die bewegte

Filmsequenz eingespielt werden kann. Das ist eine kleine Revolution. Nicht zuletzt, weil man jetzt einfach einen beliebigen Gesichtsausdruck in ein Gesichtsmodell übertragen kann.

Inzwischen haben bereits einige Produzenten bei Theobalt angeklopft. Bislang muss er sie allerdings enttäuschen. „Wir müssen unser Gesichtsmodell noch

optimieren, vor allem die Lippenbewegungen. Denn der Mensch ist enorm gut darin, kleine Ungenauigkeiten wahrzunehmen“, sagt der Forscher. Schließen sich die Lippen bei einem Laut nicht hundertprozentig korrekt, wirkt das sehr störend. „Aber ich denke, dass wir in ein paar Jahren so weit sein werden“, sagt Christian Theobalt. ◀

### AUF DEN PUNKT GEBRACHT

- Damit die Mimik und die Bewegungen animierter Figuren in Filmen, Computerspielen oder anderen Anwendungen in der virtuellen oder erweiterten Realität natürlich wirken, verwenden 3D-Grafikspezialisten bislang mit großem Aufwand erzeugte dreidimensionale Modelle von Gesichtern beziehungsweise Körpern.
- Christian Theobalt und seine Mitarbeiter entwickeln Methoden, um mit vergleichsweise geringem Rechenaufwand die Bewegungen von Gesichtern und Körpern anhand von Aufnahmen von einer oder wenigen Kameras bei beliebiger Beleuchtung und in beliebiger Umgebung zu analysieren und in Modelle umzusetzen.
- Die Forscher verwenden zu diesem Zweck unter anderem anatomische Modelle, die in ihrer Software hinterlegt sind, und Methoden des maschinellen Lernens.
- Dank des geringeren Aufwands, Bewegungen in dreidimensionale Modelle umzusetzen, werden so auch Anwendungen möglich, die bislang undenkbar waren. So können die Forscher die Mimik einer Person in Echtzeit auf das Gesicht einer anderen Person übertragen.

### GLOSSAR

**Bewegungsanalyse:** Bei dieser auch als *motion capturing* bezeichneten Technik werden die Bewegungen von Personen mit verschiedenen Methoden in drei Dimensionen erfasst. Ältere Verfahren sind dabei auf Marker und genau definierte Aufnahmebedingungen angewiesen.

**Maschinelles Lernen:** Computer werden mithilfe vieler Datensätze für unterschiedliche Aufgaben trainiert. So lernen sie, Objekte wie etwa einen Tisch zu erkennen, auch wenn sie bislang nur ähnliche Gegenstände gesehen haben oder den Gegenstand aus einer ungewohnten Perspektive sehen.



Für Forscher, Entdecker, Wissenschaftler  
- und solche, die es werden wollen:

# Junge Wissenschaft



Das einzige europäische Wissenschaftsmagazin mit begutachteten Beiträgen junger Nachwuchsforscher.

Wissenschaftliche Erstveröffentlichungen und das Neueste aus Mathematik, Informatik, Naturwissenschaft und Technik.

Nur im Abo. Viermal im Jahr News aus Forschung und Technik, Veranstaltungen, Porträts, Studien- und Berufsprofile.

## Vorteilsabo sichern!

[abo@verlag-jungewissenschaft.de](mailto:abo@verlag-jungewissenschaft.de)  
Stichwort: „Vorteilsabo“

## Leseprobe anfordern!

[leseprobe@verlag-jungewissenschaft.de](mailto:leseprobe@verlag-jungewissenschaft.de)  
oder per Fax 0211 / 74 95 64-29

**Vorteilsabo**  
nur **20,-€\***

für Schüler, Studenten, Referendare und Lehrer  
(4 Ausgaben für 20,00 EUR statt 30,00 EUR)\*  
\*zzgl. Versandkosten

[www.verlag-jungewissenschaft.de](http://www.verlag-jungewissenschaft.de)